Informatique théorique et applications

P. Massazza

Holonomic functions and their relation to linearly constrained languages

Informatique théorique et applications, tome 27, n° 2 (1993), p. 149-161.

<http://www.numdam.org/item?id=ITA_1993__27_2_149_0>

Informatique théorique et Applications/Theoretical Informatics and Applications
(vol. 27, n° 2, 1993, p. 149 à 161)

HOLONOMIC FUNCTIONS AND THEIR RELATION TO LINEARLY CONSTRAINED LANGUAGES [1] (*)

by P. MASSAZZA [2]

Communicated by C. CHOFFRUT

Abstract. – *In this paper the class of Linearly Constrained Languages* (*LCL*) *is considered. A language L belongs to LCL iff it is the set of strings of a unambiguous context-free language L' that satisfy linear constraints on the number of occurrences of symbols. We prove that every language in LCL admits a holonomic generating function, namely a function that satisfies a linear differential equation with polynomial coefficients.*

Résumé. – *Dans cet article on considère la classe des langages avec des restrictions linéaires* (*LCL*). *Un langage L est dans la classe LCL si et seulement si il est l'ensemble de mots d'un langage algébrique non ambigu L' qui vérifient des restrictions linéaires sur le nombre des occurrences des lettres. Nous montrons que à chaque langage dans LCL on peut associer une fonction génératrice holonome.*

INTRODUCTION

Generating functions are widely used in order to study properties of languages: in this context the generating function of a language L is considered to be the generating function of the sequence $\{a_n\}$ where a_n is the number of strings of L having length n. As a most famous result, a theorem by Chomsky-Schuetzenberger [5] states that every unambiguous context-free (c.f.) language admits an algebraic generating function. This fact has lead to the development of a technique that is used in certain cases to solve the ambiguity problem for c.f. languages: by showing that the generating function

(*) Received november 1991, accepted July 1992.
(1) This research was partially supported by CEC under Esprit B.R.A. Working group No. 3166 – ASMICS.
(2) Dipartimento di Scienze dell'Informazione Università degli Studi di Milano via Comelico 39, 20135 Milano, Italy.

of a c.f. language L is not algebraic, it follows that L is inherently ambiguous [8]. Generating functions are also invaluable tools for counting and decision problems for languages [3, 15, 14].

In these settings, one question raises quite naturally: is it possible to develop a taxonomy of languages w.r.t. the analytical properties of the associated generating functions? A partial answer lies in the following diagram that illustrates known results about classes of languages and classes of generating functions.

$$\begin{array}{ccc} \text{Regular languages} & \rightarrow & \text{Rational generating functions} \\ \cap & & \cap \\ \text{Unambiguous c.f. languages} & \rightarrow & \text{Algebraic generating functions} \end{array}$$

An interesting task is that of finding classes of languages and classes of generating functions to add to this diagram. As a first step in this direction, we consider the class of the holonomic functions, introduced by Bernstein [1, 2] in the early 70's and, more recently, studied by Zeilberger [19] in order to prove special functions identifies. Informally a function is said to be holonomic if it satisfies a linear differential equation with polynomial coefficients (a system of equations in the multivariate case): this class of functions is an immediate and interesting extension of the class of the algebraic functions.

Then, we consider the class LCL (Linearly Constrained Languages) that is obtained by taking the intersection of the class of unambiguous c.f. languages and the class of languages whose words satisfy a finite set of linear constraints on the number of occurrences of symbols: we prove that languages in LCL have holonomic generating functions. This result leads us to extend the previous diagram as follows:

$$\begin{array}{ccc} \text{Regular languages} & \rightarrow & \text{Rational generating functions} \\ \cap & & \cap \\ \text{Unambiguous c.f. languages} & \rightarrow & \text{Algebraic generating functions} \\ \cap & & \cap \\ \text{LCL} & \rightarrow & \text{Holonomic generating functions} \end{array}$$

Besides a general interest in the class LCL itself, a particular motivation of this study relies on the fact that counting problems for languages with holonomic generating functions can be solved efficiently [9]. Moreover, the asymptotics of linear recurrences has been widely studied [18] and the counting sequence corresponding to a holonomic generating function satisfies a linear recurrence with polynomials coefficients. As an example, given a hypercube p in an n-dimensional grid, the set of random paths that start at the origin and end up inside p is easily coded with a language in LCL. Then, for any input n the number of paths of length n can be computed in parallel (with a polynomial number of processors) in time $O(\log^2 n)$.

1. PRELIMINARIES

In this section we recall some basic notions on formal series and generating functions of languages: classical reference books are [16, 4]. We additionally point out some results about linear diophantine equations that we shall need in the sequel.

Let X be a set of n symbols, $X = \{x_1, \ldots, x_n\}$, we denote by X^C the free commutative monoid generated by X. Given a semi-ring $\mathbb{K}$ we have the following:

DEFINITION 1: *A formal series on X (in commutative variables) is a function*

$$\varphi: \quad X^C \to \mathbb{K}.$$

We think to a formal series φ as a formal power series

$$\varphi = \sum_{\underline{x}^{\underline{a}} \in X^C} \varphi(\underline{x}^{\underline{a}}) \cdot \underline{x}^{\underline{a}}$$

and we indicate by $\mathbb{K}[\![X]\!]$ the set of the formal series on X with coefficients in $\mathbb{K}$.

We consider on the set $\mathbb{K}[\![X]\!]$ the usual operations of sum, Cauchy product, Hadamard product, external product, partial derivative, together with an operator E (called *Total substitution*) that maps formal series in n variables onto formal series in one variable.

DEFINITION 2: *Let φ, ψ be two formal series in $\mathbb{K}[[X]]$, and let k be an element of $\mathbb{K}$. We define the following operations.*

- *Sum*: $(\varphi + \psi)(\underline{x}^{\underline{a}}) = \varphi(\underline{x}^{\underline{a}}) + \psi(\underline{x}^{\underline{a}})$.
- *Cauchy product*: $(\varphi \cdot \psi)(\underline{x}^{\underline{a}}) = \sum_{\underline{x}^{\underline{b}} \cdot \underline{x}^{\underline{c}} = \underline{x}^{\underline{a}}} \varphi(\underline{x}^{\underline{b}}) \psi(\underline{x}^{\underline{c}})$.
- *Hadamard product*: $(\varphi \odot \psi)(\underline{x}^{\underline{a}}) = \varphi(\underline{x}^{\underline{a}}) \psi(\underline{x}^{\underline{a}})$.
- *External product*: $(k\varphi)(\underline{x}^{\underline{a}}) = k\varphi(\underline{x}^{\underline{a}})$.
- *Partial derivative*:

$$(\partial_i \varphi)(x_1^{a_1} \ldots x_i^{a_i} \ldots x_n^{a_n}) = (a_i + 1)\, \varphi(x_1^{a_1} \ldots x_i^{a_i+1} \ldots x_n^{a_n}).$$

• *Total substitution*: $(E(\varphi))(x^k) = \sum_{a_1+\ldots+a_n=k} \varphi(x_1^{a_1}\ldots x_n^{a_n})$.

A formal series $\psi \in \mathbb{K}[[X]]$ is called *proper* if $\psi(\varepsilon)=0$. Given a vector of proper formal series $\Psi=(\psi_1, \ldots, \psi_n)$, $\psi_j \in \mathbb{K}[[Y]]$, and a series $\varphi \in \mathbb{K}[[X]]$ the *composition* $\varphi \circ \Psi \in \mathbb{K}[[Y]]$ is the series

$$\varphi \circ \Psi = \sum_{\underline{x}^a \in X^C} \varphi(\underline{x}^a) \cdot \underline{\psi}^a.$$

In the sequel we shall consider formal series having coefficients in the semiring $\mathbb{N}$, that is elements of the integral domain $[\mathbb{N}[[X]]; +, \cdot, \mathbf{0}, \mathbf{1}]$, where the unities $\mathbf{0}$, $\mathbf{1}$ are the series

$$\mathbf{0}(\underline{x}^a)=0$$

and

$$\mathbf{1}(\underline{x}^a)=\begin{cases} 1 & \text{if } \underline{x}^a=\varepsilon, \\ 0 & \text{otherwise.} \end{cases}$$

Besides the class of polynomials $\mathbb{N}[X]$, rational and algebraic formal series form two well-known subclasses of $\mathbb{N}[[X]]$, respectively denoted by $\mathbb{N}[[X]]_r$ and $\mathbb{N}[[X]]_a$. We recall that neither $\mathbb{N}[[X]]_r$ (in the multivariate case) nor $\mathbb{N}[[X]]_a$ are closed under the Hadamard product (*see* for instance [17], [13]): this is one of the reasons that lead to consider the class of the holonomic formal series, denoted by $\mathbb{N}[[X]]_h$.

DEFINITION 3: *A formal series* $\varphi \in \mathbb{N}[[X]]$ *is said to be holonomic iff there exist some polynomials*

$$p_{ij} \in \mathbb{N}[X], \qquad 1 \leqq i \leqq n, \qquad 0 \leqq j \leqq d_i,$$

such that

$$\sum_{j=0}^{d_i} p_{ij}\, \partial_i^j \varphi = 0, \qquad 1 \leqq i \leqq n.$$

Closure properties of holonomic functions are studied in [19, 12, 13]. The main result is given by the following:

THEOREM 1: *The class* $\mathbb{N}[\![X]\!]_h$ *is closed under the operations of sum, Cauchy product, Hadamard product, total substitution, right-composition with algebraic series.*

Moreover, it can be easily proved that $\mathbb{N}[[X]]_r \subset \mathbb{N}[[X]]_a \subset \mathbb{N}[[X]]_h$: so, we can consider $\mathbb{N}[[X]]_h$ as an interesting extension of $\mathbb{N}[[X]]_a$.

Given an alphabet Σ and a language $L \subseteq \Sigma^*$, the *generating function* of L is the formal power series $F_L \in \mathbb{N}[[z]]$ defined as

$$F_L = \sum_{k \geq 0} F_L(z^k) z^k$$

where $F_L(z^k) = \# \{ w \mid w \in L \wedge |w| = k \}$. This series can be interpreted as a function in the complex variable z; in particular we note that it is an analytic function at the origin with a convergence radius ρ, $1/\#\Sigma \leq \rho \leq 1$.

Let

$$a_1 x_1 + \ldots + a_n x_n = 0, \qquad a_i \in \mathbb{Z} \tag{1}$$

be a linear homogeneous diophantine equation. We denote by $\underline{x}$ a nonnegative solution of (1) ($\underline{x} \in \mathbb{N}^n$) and by S the set of all nonnegative solutions. Moreover, we consider a relation $< \subset \mathbb{N}^n \times \mathbb{N}^n$ such that

$$(r_1, \ldots, r_n) < (s_1, \ldots, s_n)$$

iff $r_i \leq s_i$ $(1 \leq i \leq n)$ and there exists an index j s.t. $r_j < s_j$ $(1 \leq j \leq n)$. A solution $\underline{x}$ of (1) is said *minimal* if $\underline{x} \neq \underline{0}$ and there exists no solution $\underline{y}$ of (1) s.t. $\underline{y} < \underline{x}$.

In [10, 11] it is proved that S is a finitely generated monoid; more precisely, it is shown that the set $S_{\min} = \{ \underline{y}_1, \ldots, \underline{y}_r \}$ of minimal solutions is a finite basis for S, that is every solution $\underline{x}$ can be univocally expressed as

$$\underline{x} = c_1 \underline{y}_1 + \ldots c_r \underline{y}_r, \qquad c_j \in \mathbb{N}.$$

Similar results hold in the nonhomogeneous case [6]. Consider the equation

$$b_1 x_1 + \ldots + b_n x_n + p = 0, \qquad b_i, p \in \mathbb{Z}. \tag{2}$$

The set of solutions of (2) is given by

$$S = \{ \underline{r} \in \mathbb{N}^n \mid \underline{r} = \underline{s} + \underline{t},\ \underline{s} \in S_{\min},\ \underline{t} \in \hat{S} \}$$

where $\hat{S}$ is the set of solutions of the associated homogeneous equation

$$b_1 x_1 + \ldots + b_n x_n = 0 \tag{3}$$

and

$$S_{\min} = \{ (r_1, \ldots, r_n) \mid (r_1, \ldots, r_n, 1) \in \hat{S}_{\min} \}$$

with $\hat{S}_{\min}$ denoting the set of minimal solutions of the equation

$$b_1 x_1 + \ldots + b_n x_n + p x_{n+1} = 0. \tag{4}$$

2. THE CLASS LCL

Let Σ be an alphabet $\Sigma = \{\sigma_1, \ldots, \sigma_n\}$. We consider the set of couples $\langle G, \mathscr{E} \rangle$ where G is an unambiguous c.f. grammar and $\mathscr{E}$ belongs to the set $\mathscr{E}V_n$ of constraints in n variables (*n-constraints*). An n-constraint is an expression that is built up of basic elements called *n-atoms*.

DEFINITION 4: n-atom) *Let $R = \{=, \neq, \leqq, \geqq, <, >\}$ be the set of symbols of relation. A n-atom is an element of the set*

$$\mathrm{At} = \mathbb{Z}^n \times \mathrm{R} \times \mathbb{Z}.$$

DEFINITION 5: (n-constraint) *Every n-atom is an n-constraint. If $\mathscr{E}_1$ and $\mathscr{E}_2$ are n-constraints then*

- $\mathscr{E}_1 \vee \mathscr{E}_2$,
- $\mathscr{E}_1 \wedge \mathscr{E}_2$,
- $\mathscr{E}_1 \Rightarrow \mathscr{E}_2$,
- $\neg \mathscr{E}_1$

are n-constraints too. No other expression apart from these is an n-constraint.

We introduce on $\mathscr{E}V_n$ a denotation function $[\![\ ]\!]$,

$$[\![\ ]\!] : \mathscr{E}V_n \mapsto 2^{\Sigma^*},$$

that associates with every expression a language called the *semantics* of the expression.

DEFINITION 6: *Let $[i_1, \ldots, i_n, r, k]$ be an n-atom and $\mathscr{E}_1$, $\mathscr{E}_2$ two elements of $\mathscr{E}V_n$. Then we have:*

- $[\![[i_1, \ldots, i_n, r, k]]\!] = \{w \in \Sigma^* \mid i_1 |w|_{\sigma_1} + \ldots + i_n |w|_{\sigma_n}\, r\, k\}$,
- $[\![\mathscr{E}_1 \vee \mathscr{E}2]\!] = [\![\mathscr{E}_1]\!] \cup [\![\mathscr{E}_2]\!]$,
- $[\![\mathscr{E}_1 \wedge \mathscr{E}_2]\!] = [\![\mathscr{E}_1]\!] \cap [\![\mathscr{E}_2]\!]$,
- $[\![\neg \mathscr{E}_1]\!] = \Sigma^* \setminus [\![\mathscr{E}_1]\!]$,
- $[\![\mathscr{E}_1 \Rightarrow \mathscr{E}_2]\!] = [\![\neg \mathscr{E}_1]\!] \cup [\![\mathscr{E}_2]\!]$.

Now, we can formally define the language denoted by a couple $\langle G, \mathscr{E} \rangle$.

DEFINITION 7: *Let G be an unambiguous c.f. grammar and $\mathscr{E}$ an n-constraint. The couple $\langle G, \mathscr{E} \rangle$ denotes the language*

$$L_{\langle G, \mathscr{E} \rangle} = L_G \cap [\![\mathscr{E}]\!].$$

DEFINITION 8: (LCL) *The class LCL (linearly constrained languages) is the class of the languages defined by couples $\langle G, \mathscr{E} \rangle$.*

Example 1: *Let us consider the language L that contains the arithmetic expressions in prefix form (built up of constants $a = +1$, $b = -1$ and the operator $+$) that have value* 0 *(note that L is not c.f.). It holds that $L = L_{\langle G, \mathscr{E} \rangle}$ where:*

$$\begin{aligned}
&\Sigma = \{+, a, b\},\\
&G = (\{S\}, \Sigma, \{S \to +SS, S \to a, S \to b\}, S),\\
&\mathscr{E} = [0, 1, -1, =, 0].
\end{aligned}$$

3. FORMAL SERIES AND THE CLASS LCL

Given a couple $\langle G, \mathscr{E} \rangle$ we consider two formal series ψ_G, $\chi_{\mathscr{E}}$, associated with the languages L_G and $[\![\mathscr{E}]\!]$.

DEFINITION 9: (ψ_L) *Let L be a language. The series $\psi_L \in \mathbb{N}[[\Sigma]]$, is the series*

$$\psi_L = \sum_{w \in L} \varphi(w),$$

where $\varphi : \Sigma^* \mapsto \Sigma^C$ is the natural morphism that associates with every word $w \in \Sigma^*$, $|w|_{\sigma_1} = i_1, \ldots, |w|_{\sigma_n} = i_n$, the monomial $\sigma_1^{i_1} \ldots \sigma_n^{i_n}$.

In the sequel, for brevity, we write φ_G instead of φ_{L_G}; we note that the coefficient of the monomial $\sigma_1^{i_1} \ldots \sigma_n^{i_n}$ in the series ψ_G is equal to the number of words in L_G associated with it by φ,

$$\psi_G(\sigma_1^{i_1} \ldots \sigma_n^{i_n}) = \#\{w \in L_G \mid |w|_{\sigma_1} = i_1, \ldots, |w|_{\sigma_n} = i_n\}.$$

DEFINITION 10: ($\chi_{\mathscr{E}}$) *Let $\mathscr{E} \in \mathscr{E}V_n$. The series $\chi_{\mathscr{E}}$ is the characteristic function of the support of the series $\psi_{[\![\mathscr{E}]\!]}$,*

$$\chi_{\mathscr{E}}(\sigma_1^{i_1} \ldots \sigma_n^{i_n}) = \begin{cases} 1 & \text{if } \exists\, w \in [\![\mathscr{E}]\!] \, (|w|_{\sigma_1} = i_1, \ldots, |w|_{\sigma_n} = i_n), \\ 0 & \text{otherwise.} \end{cases}$$

Example 2: Let $\Sigma = \{a, b, c\}$ *and consider the grammar*

$$G = (\{S\}, \Sigma, \{S \to aS, S \to bS, S \to cS, S \to \varepsilon\}, S),$$

then

$$\psi_G = \sum_{\substack{i \geqq 0 \\ j \geqq 0 \\ k \geqq 0}} \binom{i+j+k}{i, j, k} a^i b^j c^k = \frac{1}{1-(a+b+c)}.$$

Example 3: Let $\Sigma = \{a, b, c\}$. *Given the n-constraint*

$$\mathscr{E} = [1, -1, 0, =, 0] \wedge [0, 1, -1, =0]$$

it follows that

$$[\![\mathscr{E}]\!] = \{ w \in \Sigma^* \mid |w|_a = |w|_b = |w|_c \}$$

and

$$\chi_{\mathscr{E}} = \sum_{n=0}^{\infty} a^n b^n c^n = \frac{1}{1-abc}.$$

It is immediate to observe that the generating function $F_{L\langle G, \mathscr{E}\rangle}$ can be expressed in terms of the series ψ_G and $\chi_{\mathscr{E}}$ that is:

$$F_{L\langle G, \mathscr{E}\rangle} = E(\psi_G \odot \chi_{\mathscr{E}}). \tag{5}$$

In fact, the Hadamard product of ψ_G and $\chi_{\mathscr{E}}$ is a series $\theta_{L\langle G, \mathscr{E}\rangle} \in \mathbb{N}[[\Sigma]]$,

$$\theta_{L\langle G, \mathscr{E}\rangle} = \psi_G \odot \chi_{\mathscr{E}},$$

such that the coefficient in it of the monomial $\sigma_1^{i_1} \ldots \sigma_n^{i_n}$ is exactly the number of words of length $i_1 + \ldots + i_n$ that belong to $L_{\langle G, \mathscr{E}\rangle}$ and contain i_j symbols of type $\sigma_j (1 \leqq j \leqq n)$. By applying the operator E to $\theta_{L\langle G, \mathscr{E}\rangle}$ we obtain a series in $\mathbb{N}[\![z]\!]$ such that the coefficient in it of z^k is the number of words in $L_{\langle G, \mathscr{E}\rangle}$ of length k, that is

$$E(\theta_{L\langle G, \mathscr{E}\rangle}) = F_{L\langle G, \mathscr{E}\rangle}.$$

Example 4: Referring to examples 2, 3, we have

$$E(\psi_G \odot \chi_{\mathscr{E}}) = \sum_{n \geqq 0} \frac{3n!}{(n!)^3} z^{3n}.$$

4. GENERATING FUNCTIONS OF LANGUAGES IN LCL

The relation (5) leads us to consider the analytical properties of series ψ_G and $\chi_{\mathscr{E}}$. We show that these series are algebraic and rational respectively. As a consequence, the closure properties of the class $\mathbb{N}[[\Sigma]]_h$ lead us to state that every language in LCL admit a holonomic generating function.

LEMMA 1: *Let G be an unambiguous c.f. grammar. The series ψ_G turns out to be algebraic hence holonomic.*

Proof: Given an unambiguous c.f. grammar $G=(V, \Sigma, P, S)$ it is possible to find an equivalent grammar $G_1=(V_1, \Sigma, P_1, S_1)$ that is in Chomsky normal form.

We associate with each nonterminal S_i of G_1, $(1 \leqq i \leqq p)$, a formal series $\varphi_i \in \mathbb{N}[[\Sigma]]$ such that for all monomials $\sigma_1^{i_1} \ldots \sigma_n^{i_n} \in \Sigma^C$ it holds

$$\varphi_i(\sigma_1^{i_1} \ldots \sigma_n^{i_n}) = \#\{ w \in \Sigma^* \mid S_i \overset{*}{\Rightarrow} w,\ |w|_{\sigma_1} = i_1, \ldots, |w|_{\sigma_n} = i_n \}.$$

From the set of productions P_1 we obtain a system of algebraic equations

$$S_i = p_i, \qquad 1 \leqq i \leqq p,$$

where $p = \# V_1$, $p_i \in \mathbb{N}[\Sigma \vee V_1]$ and $p_i(\varepsilon)=0$, $p_i(S_j)=0$. Since this is a proper system, it admits a unique solution (*see* [16]) that is the vector $\Phi=(\varphi_1, \ldots, \varphi_p)$. Then, the first component φ_1 is an algebraic series. Hence, since $\psi_G = \varphi_1$ we conclude that ψ_G is algebraic and, a fortiori, holonomic $(\mathbb{N}[[\Sigma]]_a \subset \mathbb{N}[[\Sigma]]_h)$. ■

Now, we consider an n-constraint $\mathscr{E}$ that, w.l.o.g., can be thougth to be in disjunctive normal form,

$$\mathscr{E} = \mathscr{E}_1 \vee \ldots \vee \mathscr{E}_m,$$

where $\mathscr{E}_i$ is an n-atom or the negation of an n-atom $(1 \leqq i \leqq m)$. By the inclusion-exclusion principle we write the series $\chi_{\mathscr{E}}$ as the sum of $2^m - 1$ series

$$\chi_{\mathscr{E}} = \chi_{\mathscr{E}_1} + \ldots + \chi_{\mathscr{E}_m} - \chi_{\mathscr{E}_1 \wedge \mathscr{E}_2} - \ldots - \chi_{\mathscr{E}_1 \wedge \mathscr{E}_m} + \chi_{\mathscr{E}_1 \wedge \mathscr{E}_2 \wedge \mathscr{E}_3} + \ldots (-)^{m+1} \chi_{\mathscr{E}_1 \wedge \ldots \wedge \mathscr{E}_m}. \quad (6)$$

Then, the problem consists of studying series of the type $\chi_{\mathscr{E}'}$, where $\mathscr{E}'$ is an n-constraint that is the intersection of a finite number of n-atoms or negated n-atoms: such series are the result of the Hadamard product of the series associated with each n-atom,

$$\left.\begin{array}{c}\chi_{\mathscr{E}'}=\chi_{\mathscr{E}''_1\wedge\ldots\wedge\mathscr{E}''_p}=\chi_{\mathscr{E}''_1}\odot\ldots\odot\chi_{\mathscr{E}''_p},\\ \chi_{\mathscr{E}''_j}=\chi_{[i_1,\ldots,i_n,r,k]}\quad\text{or}\quad\chi_{\neg[i_1,\ldots,i_n,r,k]}.\end{array}\right\}\tag{7}$$

Expressions (6), (7) lead to analyze series $\chi_{\mathscr{E}}$ where $\mathscr{E}$ is an n-atom. In particular, the only case to be considered is $\mathscr{E}=[i_1,\ldots,i_n,\geqq,k]$: this is due to the following equalities (we say that two n-constraints are equal if they denote the same language),

- $[i_1,\ldots,i_n,=,k]=\neg[i_1,\ldots,i_n,\geqq,k+1]\wedge[i_1,\ldots,i_n,\geqq,k]$,
- $[i_1,\ldots,i_n,\neq,k]=\neg[i_1,\ldots,i_n,\geqq,k]\vee[i_1,\ldots,i_n,\geqq,k+1]$,
- $[i_1,\ldots,i_n,\leqq,k]=\neg[i_1,\ldots,i_n,\geqq,k+1]$,
- $[i_1,\ldots,i_n,<,k]=\neg[i_1,\ldots,i_n,\geqq,k]$,
- $[i_1,\ldots,i_n,>,k]=[i_1,\ldots,i_n,\geqq,k+1]$.

To carry out our analysis we need the following lemma.

LEMMA 2: *Let* $\mathscr{E}=[i_1,\ldots,i_n,=,k]$ ($\mathscr{E}=\neg[i_1,\ldots,i_n,=,k]$). *Then the series* $\chi_{\mathscr{E}}$ *is rational, hence it belongs to* $\mathbb{N}[[\Sigma]]_h$.

Proof: We first consider the case $\mathscr{E}=[i_1,\ldots,i_n,=,k]$.

Every solution $\underline{y}'=(y'_1,\ldots,y'_n)$ of the linear diophantine equation (in the n variables y_j)

$$i_1y_1+\ldots+i_ny_n=k,\tag{8}$$

uniquely identifies a monomial $\underline{\sigma}^{\underline{y}'}\in\Sigma^C$ such that

$$\varphi^{-1}(\underline{\sigma}^{\underline{y}'})\subseteqq[\![\mathscr{E}]\!].$$

Symmetrically, a word $w\in[\![\mathscr{E}]\!]$ univocally identifies a solution $\underline{y}''=(y''_1,\ldots,y''_n)$ of equation (8) such that

$$w\in\varphi^{-1}(\underline{\sigma}^{\underline{y}''}).$$

By the results of section 1 every solution of (8) is of the form $\underline{z}=\underline{x}+\underline{y}$ where $\underline{x}$ belongs to the set of minimal solutions of (8) $S_{\min}=\{\underline{x}_1,\ldots,\underline{x}_r\}$ and $\underline{y}$ belongs to the set of solutions of the associated homogeneous equation with basis $B=\{\underline{y},\ldots,\underline{y}_p\}$. From $S_{\min}$ and B we obtain $r+p$ monomials $\underline{\sigma}^{\underline{x}_1},\ldots,\underline{\sigma}^{\underline{x}_r},\underline{\sigma}^{\underline{y}_1},\ldots,\underline{\sigma}^{\underline{y}_p}$ and a rational series

$$\xi_{\mathscr{E}}=(\underline{\sigma}^{\underline{x}_1}+\ldots+\underline{\sigma}^{\underline{x}_r})(\underline{\sigma}^{\underline{y}_1}+\ldots+\underline{\sigma}^{\underline{y}_p})^*=\frac{(\underline{\sigma}^{\underline{x}_1}+\ldots+\underline{\sigma}^{\underline{x}_r})}{1-(\underline{\sigma}^{\underline{y}_1}+\ldots+\underline{\sigma}^{\underline{y}_p})}$$

such that

$$\xi_{\mathscr{E}}(\sigma_1^{i_1}\ldots\sigma_n^{i_n})=\begin{cases} c_{i_1,\ldots,i_n}\neq 0 & \text{if } \exists\, w\in[\![\mathscr{E}]\!]\,(|w|_{\sigma_1}=i_1,\ldots,|w|_{\sigma_n}=i_n),\\ 0 & \text{otherwise.}\end{cases}$$

It is immediate to see that coefficients $c_{i_1,\ldots,i_n}$ can be greater than 1, since, given a monomial $\sigma_1^{j_1}\ldots\sigma_n^{j_n}$, it might be

$$\sigma_1^{j_1}\ldots\sigma_n^{j_n}=\underline{\sigma}^{\underline{x}_i}(\underline{\sigma}^{\underline{y}_1})^{h_1}\ldots(\underline{\sigma}^{\underline{y}_r})^{h_r}=\underline{\sigma}^{\underline{x}_l}(\underline{\sigma}^{\underline{y}_1})^{k_1}\ldots(\underline{\sigma}^{\underline{y}_r})^{k_r}$$

with $\underline{h}\neq\underline{k}$ and $i\neq l$. The series $\chi_{\mathscr{E}}$ turns out to be the characteristic series of the support of $\xi_{\mathscr{E}}$. So, we are faced with the problem of determining an unambiguous commutative rational expression that denotes the same language of $(\underline{\sigma}^{\underline{x}_1}+\ldots+\underline{\sigma}^{\underline{x}_r})(\underline{\sigma}^{\underline{y}_1}+\ldots+\underline{\sigma}^{\underline{y}_p})^*$. Since in a commutative monoid every rational set is unambiguous rational [7] we conclude that $\chi_{\mathscr{E}}$ is rational, hence holonomic.

In the case $\mathscr{E}=\neg\,[i_1,\ldots,i_n,=,k]$ the series $\chi_{\mathscr{E}}$ is rational too, since it can be expressed as the difference of rational series

$$\chi_{\mathscr{E}}=\sum_{\underline{j}\neq\underline{0}}1\,\underline{\sigma}^{\underline{j}}-\chi_{[i_1,\ldots,i_n,=,k]}$$

$$=\left(\frac{1}{1-\sigma_1}-1\right)\ldots\left(\frac{1}{1-\sigma_n}-1\right)-\chi_{[i_1,\ldots,i_n,=,k]}.\quad\blacksquare$$

Now, let's turn to the case $\mathscr{E}=[i_1,\ldots,i_n,\geqq,k]$ and prove the following

LEMMA 3: *Let*

$$[i_1,\ldots,i_n,\geqq,k]$$

be an n-atom. The series $\chi_{[i_1,\ldots,i_n,\geqq,k]}$ turns out to be holonomic,

$$\chi_{[i_1,\ldots,i_n,\geqq,k]}\in\mathbb{N}\,[[\Sigma]]_h.$$

Proof: The language $[\![[i_1,\ldots,i_n,\geqq,k]]\!]$ is the set of words w s.t.

$$i_1|w|_{\sigma_1}+\ldots+i_n|w|_{\sigma_n}\geqq k.$$

Let us consider a symbol $\tau\notin\Sigma$ and let $l=i_1|w|_{\sigma_1}+\ldots+i_n|w|_{\sigma_n}-k$. By inserting (in all the possible ways) in each word w l occurrences of the symbol τ we get words v that satisfy the equation

$$i_1|v|_{\sigma_1}+\ldots+i_n|v|_{\sigma_n}-|v|_\tau=k,$$

and belong to the language (on $\Sigma \cup \{\tau\}$)

$$[\![[i_1, \ldots, i_n, -1, =, k]]\!].$$

At last, the series $\chi_{[i_1, \ldots, i_n, \geqq, k]}$ is obtained by taking the right composition of $\chi_{[i_1, \ldots, i_n, -1, =, k]}$, with the vector of monomials $\Psi = (\sigma_1, \ldots, \sigma_n, \mathbf{1})$,

$$\chi_{[i_1, \ldots, i_n, \geqq, k]} = \chi_{[i_1, \ldots, i_n, -1, =, k]} \circ \Psi.$$

Informally this operation corresponds to assign weight 0 to the symbol τ. By theorem 1 we conclude that $\chi_{[i_1, \ldots, i_n, \geqq, k]} \in \mathbb{N}[[\Sigma]]_h$. ■

Remark: Since in $[\![[i_1, \ldots, i_n, -1, =, k]]\!]$ there are no two different words w, w' s.t.

$$|w|_{\sigma_1} = |w'|_{\sigma_1}, \ldots, |w|_{\sigma_n} = |w'|_{\sigma_n}, |w|_\tau \neq |w'|_\tau,$$

the series $\chi_{[i_1, \ldots, i_n, \geqq, k]}$ that we obtain is indeed the characteristic series of the support of $\psi_{[i_1, \ldots, i_n, \geqq, k]}$.

Moreover, the composition $\chi_{[i_1, \ldots, i_n, -1, =, k]} \circ \Psi$ is well defined even if Ψ is not a vector of proper series. In fact for each n-tuple $i_1, \ldots, i_n$ the set T of monomials of the form $\sigma_1^{i_1} \ldots \sigma_n^{i_n} \tau^j$ such that

$$\chi_{[i_1, \ldots, i_n, -1, =, k]}(\sigma_1^{i_1} \ldots \sigma_n^{i_n} \tau^j) \neq 0$$

is finite (more precisely $\# T = 1$).

We are now ready to prove the main result concerning the class LCL.

THEOREM 2: *Every language in the class LCL admits a holonomic generating function.*

Proof: By relation (5) the generating function F_L of a language $L = L_{\langle G, \mathscr{E} \rangle}$, can be expressed as $F_L = E(\psi_G \odot \chi_{\mathscr{E}})$. We know that $\chi_{\mathscr{E}}$ is the sum of series that are the Hadamard product of holonomic series (equalities (6), (7) and lemma 3): by recalling the closure properties of $\mathbb{N}[[\Sigma]]_h$ (theorem 1) it follows that $\chi_{\mathscr{E}}$ is holonomic. Hence, since ψ_G is holonomic (lemma 1) we conclude that F_L is holonomic too (theorem 1 again). ■

ACKNOWLEDGEMENTS

I wish to thank Alberto Bertoni for many stimulating conversations and useful suggestions.

REFERENCES

1. I. N. BERNSTEIN, Modules over a ring of differential operators. A study of the fundamental solutions of equations with constant coefficients, *Functional Anal. Appl.*, *5*, (2), 1971, p. 1-16 (in Russian), p. 89-101 (English translation).
2. I. N. BERNSTEIN, The analytic continuation of generalized functions with respect to a parameter, *Functional Anal. Appl.*, *6*, (4), 1972, p. 26-40 (in Russian), p. 273-285 (English translation).
3. A. BERTONI, M. GOLDWURM and P. MASSAZZA, Counting problems and algebraic formal power series in noncommuting variables, *Inform. Process. Lett.*, *34*, 1990, p. 117-121.
4. J. BERSTEL and C. REUTENAUER, Rational series and their languages, Springer-Verlag, Berlin Heidelberg, 1988.
5. N. CHOMSKY and M. P. SCHUETZENBERGER, The algebraic theory of context-free languages, *Computer Programming and Formal Systems*, North-Holland, Amsterdam, 1963, p. 118-161.
6. M. CLAUSEN and A. FORTENBACHER, Efficient solution of linear diophantine equations, *J. Symbolic Comput.*, *8*, 1989, p. 201-216.
7. S. EILENBERG and M. P. SCHUETZENBERGER, Rational sets in commutative monoids, *J. Algebra*, *13*, (2), 1969, p. 173-191.
8. P. FLAJOLET, Analytic models and ambiguity of context-free languages, *Theoret. Comput. Sci.*, *49*, 1987, p. 283-309.
9. M. GOLDWURM and P. MASSAZZA, On computing the coefficients of holonomic and algebraic multivariate formal series, *Internal Report*, *Dip. di Scienze dell'Informazione*, Univ. degli Studi di Milano, 1992.
10. G. HUET, An algorithm to generate the basis of solutions to homogeneous linear diophantine equations, *Inform. Process. Lett.*, *7*, 1978, p. 144-147.
11. J. L. LAMBERT, Une borne pour les générateurs des solutions entières positives d'une équation diophantienne linéaire, *C.R. Acad. Sci. Paris*, t. 305, série I, 1987, p. 39-40.
12. L. LIPSHITZ, D-Finite Power Series, *J. Algebra*, *122*, 1989, p. 353-373.
13. P. MASSAZZA, Problemi di conteggio e funzioni generatrici olonomiche, *Tesi di Dottorato*, *Dip. di Scienze dell'Informazione*, Univ. degli Studi di Milano, 1990.
14. P. MASSAZZA and N. SABADINI, Some applications and techniques for generating functions, *Proc. CAAP*, *LNCS*, *351*, Springer-Verlag, 1989, p. 321-336.
15. P. MASSAZZA and N. SABADINI, Holonomic generating functions and context free languages, *Proc. of the first Italian conference on algorithms and complexity*, World Scientific, Singapore, 1990, p. 148-158. Extended version to appear in: *International Journal of Foundations of Computer Science*.
16. A. SALOMAA and M. SOITTOLA, Automata-theoretic aspects of formal power series, Springer-Verlag, New York, 1978.
17. R. P. STANLEY, Differentiably finite power series, *European J. Combin.*, *1*, 1980, p. 175-188.
18. J. WIMP and D. ZEILBERGER, Resurrecting the Asymptotics of Linear Recurrences, *J. Math. Anal. Appl.*, *111*, 1985, p. 162-176.
19. D. ZEILBERGER, *A holonomic systems approach to special functions identities*, *J. Comput. Appl. Math.*, *32*, 1990, p. 321-368.